

SUPPLEMENTAL McNAMARA DECLARATION
EXHIBIT 8

Durie Tangri

Joseph C. Gratz
415-376-6407 (direct)
415-362-6666 (main)
jgratz@durietangri.com

August 9, 2021

VIA ECF

Hon. John G. Koeltl
U.S. District Judge
Daniel Patrick Moynihan
United States Courthouse
Courtroom 14A
500 Pearl St.
New York, NY 10007-1312

Re: *Hachette Book Group, Inc. et al. v. Internet Archive*
Case No. 1:20-CV-04160-JGK

Your Honor:

Pursuant to Local Civil Rule 37.2, Defendant Internet Archive respectfully requests a pre-motion discovery conference regarding a motion to compel the production of information regarding the commercial performance of books published by Plaintiffs.

In the above-captioned lawsuit, Plaintiffs contend that the Internet Archive infringed Plaintiffs' copyrights by the non-profit digital lending of library books. The Internet Archive maintains that the challenged lending constitutes fair use under 17 U.S.C. § 107.

In considering fair use, one factor courts consider is "the effect of the use upon the potential market for or value of the copyrighted work." Plaintiffs claim that the Internet Archive's digital library lending has a negative effect on the market for or value of the works. The Internet Archive disagrees, and wishes to bring forward evidence showing that lending had little or no effect on the commercial performance of the books being lent, compared to books that were not lent. "Since fair use is an affirmative defense, its proponent would have difficulty carrying the burden of demonstrating fair use without favorable evidence about relevant markets." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994). The Internet Archive submits this letter-motion in order to gather that evidence.

Specifically, in order to show that lending had little or no effect on commercial performance, the Internet Archive wishes to compare the commercial performance of books that were available for digital lending with books that were not available for digital lending. The Internet Archive has requested data

Hon. John G. Koeltl
 August 9, 2021
 Page 2

about the commercial performance¹ of all of Plaintiffs' books, broken down by month, since 2011. Plaintiffs have provided some commercial performance data, but (1) Plaintiffs have refused to provide data about books other than the works-in-suit, and (2) the data they have provided is not broken down by month. For reasons explained below, Plaintiffs can and should produce this data, and the Internet Archive asks that this Court compel them to do so.

A. Commercial performance data about books other than the works-in-suit is necessary to understand whether lending had any effect on commercial performance.

In order to argue that the challenged library lending practice did not affect commercial performance, one needs commercial performance data not only for the books that were lent out, but also of other books that were not loaned. Without that data, the Internet Archive has nothing to compare. *See, e.g., Viacom Int'l Inc. v. Youtube Inc.*, 253 F.R.D. 256, 261 (S.D.N.Y. 2008) (granting motion to compel production of data about all videos not accused of infringement because such data was necessary "to compare the attractiveness of allegedly infringing videos with that of non-infringing videos.").

Plaintiffs object that data about other books would be irrelevant. They have argued that, because there are too many other factors that could affect their commercial performance, the data won't show whether the Internet Archive's digital library lending affected commercial performance. We have no doubt that Plaintiffs will press that line of argument in cross-examination of the Internet Archive's witnesses, but such estimates are a necessary part of litigation about alleged copyright infringement. *See, e.g., Davis v. The Gap, Inc.*, 246 F.3d 152, 166–67 (2d Cir. 2001) ("Many of the accepted methods of calculating copyright damages require the court to make uncertain estimates in the realm contrary to fact.").

Plaintiffs' line of argument provides no justification for withholding the data necessary to perform the comparison in the first place.

Plaintiffs also object that producing data about *all* other books would be unduly burdensome, because there are only 127 works-in-suit. The Internet Archive agrees that the comparison does not require each and every book; instead, it requires, for each work-in-suit, one or more comparable books that were not available for digital lending at the same time as that work-in-suit. One way to identify potentially comparable books would be to identify books whose commercial performance was very similar before either book was made available for digital lending. But Plaintiffs, who are in possession of the data one would need to do that analysis, have declined to identify books they regard as comparable—because, as discussed above, they take the position that no book is comparable to any other book. Given this refusal, Plaintiffs must produce data about all books, so that the Internet Archive can identify books it

¹The types of commercial performance data the Internet Archive seeks are: the number of physical copies embodying works by distribution channel, the prices for those physical copies, the number of ebooks embodying works by distribution channel, the number of ebook loans divided by distribution channel, prices for any transaction related to ebooks, income from sales of physical books by distribution channel, and income from ebook transactions by distribution channel. RFPs 20-22, 60-67.

Hon. John G. Koeltl
August 9, 2021
Page 3

regards as comparable, and the parties can then debate, on a level playing field, whether such books are or are not comparable.

Nor is there any particular burden in retrieving the information requested. This is commercial data stored in databases, indexed by book. Plaintiffs were able to provide data about the works-in-suit by accessing such databases; this motion simply seeks the result of querying the same systems for a larger set of books.

B. Plaintiffs must produce *monthly* commercial performance data, not just annual data.

Hachette, Penguin Random House, HarperCollins, and Wiley have produced annual commercial performance data for the works-in-suit. They do not contend that monthly data does not exist; they contend only that producing monthly data would be more difficult, so they have not produced it. Indeed, Hachette produced monthly data for 2020, but not for the other requested years. And, of course, as discussed above, Plaintiffs have not produced any data for books that are not works-in-suit.

Monthly data is necessary because each book began to be available for digital lending on a particular known date, so it makes sense to compare the commercial performance *before* that particular date with the commercial performance *after* that date. Purely annual data does not provide enough detail, particularly because sales of a particular book change so drastically within a year. (*The New York Times* publishes its bestseller lists weekly, after all, for this reason.)

Monthly data is also necessary because one issue in the case is whether commercial performance was affected differently during the National Emergency Library. This was the approximately three-month period, during the first wave of shelter-in-place orders that shuttered schools and libraries, when the Internet Archive made books available for lending without imposing the strict one-to-one “paper-copies-owned-to-digital-copies-loaned” ratio it imposes during normal operations. Without monthly data, there is no way to tell whether that short-lived emergency practice affected commercial performance.

For the reasons explained above, and pursuant to Local Rule 37.2 and Your Honor’s Individual Practices Paragraphs 1F and 2B, Defendant respectfully requests a pre-motion conference with the Court to discuss: Plaintiffs’ production of monthly commercial performance data since 2011 for all of their books in print during that period.


Joseph C. Gratz

JCG:co



21st Floor
1251 Avenue of the Americas
New York, NY 10020-1104

Elizabeth A. McNamara
(212) 489-8230 tel
(212) 489-8340 fax

lizmcnamara@dwt.com

August 12, 2021

VIA ECF

Hon Ona T. Wang
United States District Court
500 Pearl Street
New York, NY 10007
Wang_NYSDChambers@nysd.uscourts.gov

Re: *Hachette Book Group, Inc. et al. v. Internet Archive*, 1:20-CV-04160-JGK

Dear Magistrate Judge Wang:

We represent plaintiffs Hachette Book Group, Inc, HarperCollins Publishers LLC, John Wiley & Sons, Inc. and Penguin Random House LLC (“Plaintiffs”). Plaintiffs respectfully submit this response to the August 9, 2021 letter motion filed by Internet Archive (“IA”).

This lawsuit involves IA’s industrial-scale copying and online distribution of Plaintiffs’ and others’ in-copyright books. IA operates an illegal ebook distribution service that threatens to destroy the legal library ebook market and otherwise harm consumer book sales if left unchecked. IA accumulates millions of hard copies of copyrighted books as cheaply as possible, scans them, and distributes digital scans of the entire book without a license to anyone in the world who signs up to IA’s website. IA argues that this is legal fair use under a manufactured theory called “controlled digital lending.”

While Plaintiffs understand that IA has copied and distributed thousands of their books without authorization, this copyright infringement action identifies a representative subset of 127 books that are all currently licensed or sold as ebooks (the “Works in Suit”), ranging from classics by Zora Neale Hurston and J.D. Salinger to contemporary thrillers and romances.

Plaintiffs have produced a vast wealth of detailed sales and related financial data concerning the Works in Suit, totaling over 670,000 rows of data in Excel. Now, after the close of document discovery and on the eve of depositions, IA seeks to compel the production of “commercial performance data,” broken down by month, distribution channel, price and income, for *all* other books published by the Plaintiffs since 2011 – an undertaking that would involve a massive amount of data concerning more than 500,000 titles.

And IA makes this extraordinary demand in order to rifle through an enormous reservoir of highly proprietary data concerning books that are *not* the Works in Suit, all in an effort to somehow select “for each work in suit, one or more comparable books that were not available for digital lending” on IA’s system. Dkt. 47, 2. In other words, because the significant financial data already provided concerning the Works in Suit apparently does not support IA’s theory on

DWT.COM

Anchorage | Bellevue | Los Angeles | New York
Portland | San Francisco | Seattle | Washington, D.C.

Hon. Ona T. Wang
August 12, 2021
Page 2

market harm, IA wants access to millions of data points concerning Plaintiffs' entire book catalogues. IA argues it is entitled to do this in order to see if any evidence might exist to support the inherently incredible theory that copying entire books and distributing them to any member of the public worldwide upon demand does not compete with Plaintiffs' sales of the same books. Even worse, IA's quest rests on the palpably false theory that it can quantify the harm caused by its infringement by comparing the sales of completely different books. Books are not interchangeable widgets and marketplace performance is driven by countless indeterminate and changing facts.

In short, Defendant's letter motion should be denied because the enormous and costly burden to Plaintiffs far outweighs the negligible value (if any) of the evidence sought, especially given the lack of legally relevant results that it will yield and the delays it will cause to the case.

THE DATA IA SEEKS IS NOT MATERIAL AND CREATES A MASSIVE BURDEN

Plaintiffs' books are distributed to millions of readers around the world through many legal channels, including a thriving market for library ebook lending. Plaintiffs sell or license ebooks to library aggregators, who provide ebook platforms that millions of library patrons use to borrow ebooks on their devices for free, instantaneously and lawfully, from any location, subject to certain limitations.

The fourth factor of the fair use analysis examines the potential market harm from Defendant's action. IA argues that it needs sales data or comparables "to argue that the challenged library lending practice did not affect commercial performance" of the Works in Suit. This argument wrongfully cabins and misinterprets the fourth factor of the fair use analysis and evidences a false assumption about what drives book sales.

The fourth factor looks beyond lost sales and focuses instead on "the effect of the [infringing use] upon the *potential* market for or value of the copyrighted work." 17 U.S.C. § 107(4). "Critically, it requires consideration of 'not only ... the market harm caused by the particular actions of the alleged infringer,' but also the market harm that would result from 'unrestricted and widespread conduct of the [same] sort.'" *Fox News Network v. TVEyes, Inc.*, 883 F.3d 169, 179-80 (2d Cir. 2018) (quoting *Campbell v. Acuff-Rose*, 510 U.S. 569, 590 (1994)).

Here, IA self-evidently harms Plaintiffs' existing markets by refusing to pay the customary fees that aggregators pay to distribute the same ebooks to library borrowers. The market harm compounds as others engage in the activity, destroying the vibrant market for licensed library ebook lending that currently exists. None of the requested data, or so-called "comparables," is necessary to or even related to this category of market harm.

Nor is the demanded sales data necessary to demonstrate potential market harm by showing that the "secondary use competes in the rightsholder's market as an effective substitute for the original." *Capital Records v. ReDigi Inc.*, 910 F.3d 649, 662 (2d Cir. 2018). Over and over the Second Circuit has recognized that a non-transformative use of a work – as here – is likely to be "an effective substitute for the original." *Id.* at 662-63; *Authors Guild v. Google Inc.*,

Hon. Ona T. Wang
August 12, 2021
Page 3

804 F.3d 202, 223, 225-26 (2d Cir. 2015) (a book publisher would have a “strong” claim of market harm if their “claim were based on Google’s converting their books into a digitized form and making that digitized version accessible to the public”).

Even if sales data is the yardstick IA elects to focus on, to the extent that inquiry is relevant, Plaintiffs have already produced hugely detailed sales data relating to the Works in Suit for many years before and during IA’s infringement.¹ Seeking an additional “comparable” set of sales data for half of million other books is not only burdensome in the extreme, it is irrelevant.

First, IA’s argument that there is no burden in producing the data is preposterous. Contrary to IA’s suggestion that this information is available at the flip of a switch, the request would require Plaintiffs to pull masses of data from multiple databases, stitch it together and figure out a way to transfer enormous datasets in a way that they do not do in the ordinary course of business. An analyst for one Plaintiff estimated that it would take his entire team several weeks or months working full time to compile the requested information – which would amount to hundreds of millions, if not billions, of rows of data. The demand seeks access to data in a manner that lies far outside the bounds of how data is accessed, amassed or referenced by the Plaintiff publishers in the ordinary course of business and would involve incalculable employee time and significant resources. IA offers no precedent where a similar demand was allowed.²

Second, since books are not fungible widgets, the entire project rests on a false premise. There is no such thing as a “comparable book” – even if “comparable” is defined as some undefined period of sales data. Should *Catcher in the Rye* have similar sales to a best-selling cookbook, no one could plausibly contend the two works were “comparable.” Moreover, sales vary dramatically over different periods of time, depending on countless events (e.g., an author dies; the book is made into a movie; it receives a great review or is selected by a state for its curriculum). No control group can begin to equate sales data.

In short, it is impossible to calculate the market harm of IA’s infringement based on the crude comparison IA proposes because there are innumerable reasons why one book sells more copies than another that have nothing to do with IA’s infringement. IA’s demand for this massive data rests on no coherent foundation. It should be denied.

Respectfully submitted,

/s/ Elizabeth A. McNamara

¹ IA demands monthly sales data for the Works in Suit. Plaintiffs have already produced monthly sales data from Hachette for 2020 and HarperCollins for a number of the works; the rest of the data reflects annual sales. To produce additional monthly sales is both burdensome and not necessary.

² IA cites the inapposite *Viacom Int’l Inc. v. YouTube Inc.*, 253 F.R.D. 256 (S.D.N.Y. 2008). Despite IA’s misleading parenthetical, that court’s order compelling YouTube, the *defendant*, to produce a minimal quantity of data concerning certain non-infringing videos to support *plaintiffs’* infringement is irrelevant to IA’s demand for billions of rows of data for every work published by Plaintiffs in the last decade.

Hon. Ona T. Wang
August 12, 2021
Page 4

cc: Counsel of Record (via ECF)



21st Floor
1251 Avenue of the Americas
New York, NY 10020-1104

Elizabeth A. McNamara
(212) 603-6437 tel
(212) 379-5237 fax

lizmcnamara@dwt.com

January 14, 2022

Hon. Ona T. Wang
United States District Court
Southern District of New York
500 Pearl Street
New York, N.Y. 10007-1312

Re: *Hachette Book Group, Inc. et al. v. Internet Archive et al.*,
1:20-CV-04160-JGK (S.D.N.Y.)

Dear Judge Wang:

Pursuant to this Court's order at the December 2, 2021 conference (Dkt. 62), we write on behalf of all parties to provide the Court with a status report concerning the discovery disputes before the Court. At the December 2 conference, three letter motions were before the Court: Defendant's August 9, 2021 letter motion (Dkt. 47), Defendant's October 29, 2021 letter motion (Dkt. 54), and Plaintiffs' November 19, 2021 letter motion (Dkt. 58) (together, the "Letter Motions").

We are pleased to inform the Court that, subject to meaningful compliance with the agreements reached, which are memorialized in a separate document, the parties have resolved the various issues identified in the Letter Motions, together with related discovery disputes. More specifically, the parties have agreed that in lieu of additional depositions, certain witnesses will provide declarations addressing identified issues/questions, and have agreed that certain additional documents will be produced (if available). Further, the parties have agreed that particular non-party depositions (and related document discovery) will occur at mutually agreed times outside the close of fact discovery pursuant to the Scheduling Order.

Finally, the parties have mutually agreed on the attached proposed revisions to the current scheduling order dated September 8, 2021 (Dkt. 51), as amended by this Court at the December 2, 2021 conference (Dkt. 60). Because the Court extended the fact discovery deadline, it becomes necessary to also grant an extension of other case deadlines to ensure that the parties have sufficient time to complete expert discovery and summary judgment briefing.^[1] The parties

^[1] Pursuant to this Court's Individual Rule I.e., the parties state that they requested two amendments of the case deadlines in this action. These requests were both granted. *See* Dkt. 44, 51.

Hon. Ona T. Wang
January 14, 2022
Page 2

also have adjusted the schedule to address the anticipated nature of expert testimony and briefing. We attach as Exhibit A the current scheduling order. We attach as Exhibit B the proposed dates for the Second Revised Scheduling Order. All parties respectfully request that the Court so-order this new proposed schedule.

Thank you for your consideration of this request.

Respectfully submitted,
Davis Wright Tremaine LLP
/s/ Elizabeth A. McNamara
Elizabeth A. McNamara

cc: Counsel of Record (via ECF)

EXHIBIT A

**IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

HACHETTE BOOK GROUP, INC.,
HARPERCOLLINS PUBLISHERS LLC,
JOHN WILEY & SONS, INC., and
PENGUIN RANDOM HOUSE LLC

Case No. 1:20-CV-04160-JGK

Plaintiffs,

v.

INTERNET ARCHIVE and DOES 1 through
5, inclusive

Defendants.

[PROPOSED] REVISED SCHEDULING ORDER

After reviewing the Letter Motion dated September 7, 2021 and for good cause shown, the Court hereby orders that the following schedule will govern this civil action:

Deadline to complete fact depositions (close of fact discovery)	Friday December 17, 2021
Deadline to complete expert reports: issues on which each party bears the burden of proof	Thursday January 14, 2022
Deadline to serve rebuttal expert reports: issues on which each party does not bear the burden of proof, but not limited to responding to the opening reports. Any response to the opening reports must be contained in the rebuttal reports, however.	Friday February 25, 2022
Deadline to serve reply reports. Reply reports are limited to responding to the rebuttal reports.	Friday March 25, 2022
End of all discovery	Friday April 1, 2022

Deadline for dispositive motions	Friday April 29, 2022
Briefs in opposition to dispositive motions due	Friday June 10, 2022
Reply briefs in further support of dispositive motions due	Friday July 8, 2022

SO ORDERED.

Dated: 9/8/21, 2021



Honorable John G. Koeltl
United States District Judge

EXHIBIT B

**IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

HACHETTE BOOK GROUP, INC.,
HARPERCOLLINS PUBLISHERS LLC,
JOHN WILEY & SONS, INC., and
PENGUIN RANDOM HOUSE LLC

Plaintiffs,

v.

INTERNET ARCHIVE and DOES 1 through
5, inclusive

Defendants.

Case No. 1:20-CV-04160-JGK

[PROPOSED] REVISED SCHEDULING ORDER

After reviewing the Joint Letter dated January 14, 2022 and for good cause shown, the Court hereby orders that the following schedule will govern this civil action:

Deadline to complete fact depositions (close of fact discovery).	January 31, 2022
Deadline to complete expert reports on issues on which each party bears the burden of proof (the “Opening Reports”).	February 25, 2022
Deadline to complete expert reports on issues on which each party does not bear the burden of proof, but not limited to responding to the opening reports (the “Rebuttal Reports”). Any response to the Opening Reports must be contained in the Rebuttal Reports, however.	April 29, 2022
Deadline to serve reply reports to the Rebuttal Reports. Reply reports are limited to responding to the Rebuttal Reports.	May 27, 2022
End of all discovery, including expert depositions.	June 10, 2022
Deadline for dispositive motions.	July 7, 2022
Filings in opposition to dispositive motions due.	September 2, 2022

Reply filings in further support of dispositive motions due.	October 7, 2022
--	-----------------

SO ORDERED.

Dated: _____, 2022

Honorable Ona T. Wang
United States Magistrate Judge